PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/01, 15/11, 15/12, 15/33, 15/62, 15/70, C07K 14/00, 14/005, 14/435, 19/00, A61K 38/16, 38/17, C12Q 1/02, 1/68, 1/70

A1

(11) International Publication Number: **WO 95/19431**

(43) International Publication Date: 20 July 1995 (20.07.95)

(21) International Application Number: PCT/US95/00829

(22) International Filing Date: 18 January 1995 (18.01.95)

(30) Priority Data:
08/183,119 18 January 1994 (18.01.94) US
08/312,604 28 September 1994 (28.09.94) US

(60) Parent Application or Grant
(63) Related by Continuation
US 08/312,604 (CIP)
Filed on 28 September 1994 (28.09.94)

(71) Applicant *(for all designated States except US)*: THE SCRIPPS RESEARCH INSTITUTE [US/US]; 10666 North Torrey Pines Road, La Jolla, CA 92037 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: BARBAS, Carlos, F., III [US/US]; 7081 Weller Street, San Diego, CA 92122 (US). GOTTESFELD, Joel, M. [US/US]; 14269 Mango Drive, San Diego, CA 92122 (US). WRIGHT, Peter, E. [US/US]; 7221 Rue Michael, La Jolla, CA 92037 (US).

(74) Agents: HAILE, Lisa, A. et al.; Spensley, Horn, Jubas & Lubitz, 5th floor, 1880 Century Park East, Los Angeles, CA 90067 (US).

(81) Designated States: AU, CA, FI, JP, NO, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: ZINC FINGER PROTEIN DERIVATIVES AND METHODS THEREFOR

(57) Abstract

An assay is described which allows for identification of novel modulating zinc finger-nucleotide binding polypeptides. Such proteins are useful for inhibiting, activating or enhancing gene expression from a zinc finger-nucleotide binding motif containing promoter or other transcriptional control element, as well as a structural gene or RNA sequence. Also described are novel zinc finger-nucleotide binding polypeptides.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

ZINC FINGER PROTEIN DERIVATIVES AND METHODS THEREFOR

This application is a continuation-in-part of application Serial No. 08/312,604, filed September 28, 1994, which is a continuation-in-part of application Serial No. 08/183,119, filed January 18, 1994.

BACKGROUND OF THE INVENTION

1. *Field of the Invention*

This invention relates generally to the field of regulation of gene expression and specifically to methods of modulating gene expression by utilizing polypeptides derived from zinc finger-nucleotide binding proteins.

2. *Description of Related Art*

Transcriptional regulation is primarily achieved by the sequence-specific binding of proteins to DNA and RNA. Of the known protein motifs involved in the sequence specific recognition of DNA, the zinc finger protein is unique in its modular nature. To date, zinc finger proteins have been identified which contain between 2 and 37 modules. More than two hundred proteins, many of them transcription factors, have been shown to possess zinc fingers domains. Zinc fingers connect transcription factors to their target genes mainly by binding to specific sequences of DNA base pairs - the "rungs" in the DNA "ladder".

Zinc finger modules are approximately 30 amino acid-long motifs found in a wide variety of transcription regulatory proteins in eukaryotic organisms. As the name implies, this nucleic acid binding protein domain is folded around a zinc ion. The zinc finger domain was first recognized in the transcription factor TFIIIA from *Xenopus* oocytes (Miller, *et al., EMBO,* 4:1609-1614, 1985; Brown, *et al., FEBS Lett.,* 186:271-274, 1985). This protein consists of nine imperfect repeats of a consensus sequence:

(Tyr, Phe)-X-Cys-X_{2-4}-Cys-X_3-Phe-X_5-Leu-X_2-His-X_{3-4}-His-X_{2-6} (SEQ ID NO: 1) where X is any amino acid.

Like TFIIIA, most zinc finger proteins have conserved cysteine and histidine residues that tetrahedrally-coordinate the single zinc atom in each finger domain. The structure of individual zinc finger peptides of this type (containing two cysteines and two histidines) such as those found in the yeast protein ADR1, the human male associated protein ZFY, the HIV enhancer protein and the *Xenopus* protein Xfin have been solved by high resolution NMR methods (Kochoyan, *et al., Biochemistry,* 30:3371-3386, 1991; Omichinski, *et al., Biochemistry,* 29:9324-9334, 1990; Lee, *et al., Science,* 245:635-637, 1989) and detailed models for the interaction of zinc fingers and DNA have been proposed (Berg, 1988; Berg, 1990; Churchill, *et al.,* 1990). Moreover, the structure of a three finger polypeptide-DNA complex derived from the mouse immediate early protein zif268 (also known as Krox-24) has been solved by x-ray crystallography (Pavletich and Pabo, *Science,* 252:809-817, 1991). Each finger contains an antiparallel β-turn, a finger tip region and a short amphipathic α-helix which, in the case of zif268 zinc fingers, binds in the major groove of DNA. In addition, the conserved hydrophobic amino acids and zinc coordination by the cysteine and histidine residues stabilize the structure of the individual finger domain.

While the prototype zinc finger protein TFIIIA contains an array of nine zinc fingers which binds a 43 bp sequence within the 5S RNA genes, regulatory proteins of the zif268 class (Krox-20, Sp1, for example) contain only three zinc fingers within a much larger polypeptide. The three zinc fingers of zif268 each recognize a 3 bp subsite within a 9 bp recognition sequence. Most of the DNA contacts made by zif268 are with phosphates and with guanine residues on one DNA strand in the major groove of the DNA helix. In contrast, the mechanism of TFIIIA binding to DNA is more complex. The amino-terminal 3 zinc fingers recognize a 13 bp sequence and bind in the major groove. Similar to zif268, these fingers also make guanine contacts primarily on one strand of the DNA. Unlike the zif268 class of proteins, zinc fingers 4 and 6 of TFIIIA each bind either in or

across the minor groove, bringing fingers 5 and 7 through 9 back into contact with the major groove (Clemens, *et al., Proc. Natl. Acad. Sci. USA*, 89:10822-10826, 1992).

The crystal structure of zif268, indicates that specific histidine (non-zinc coordinating his residues) and arginine residues on the surface of the α-helix participate in DNA recognition. Specifically, the charged amino acids immediately preceding the α-helix and at helix positions 2, 3, and 6 (immediately preceding the conserved histidine) participate in hydrogen bonding to DNA guanines. Similar to finger 2 of the regulatory protein Krox-20 and fingers 1 and 3 of Sp1, finger 2 of TFIIIA contains histidine and arginine residues at these DNA contact positions; further, each of these zinc fingers minimally recognizes the sequence GGG. Finger swap experiments between transcription factor Sp1 and Krox-20 have confirmed the 3-bp zinc finger recognition code for this class of finger proteins (Nardelli, *et al., Nature*, 349:175-178, 1989). Mutagenesis experiments have also shown the importance of these amino acids in specifying DNA recognition. It would be desirable to ascertain a simple code which specifies zinc finger-nucleotide recognition. If such a code could be deciphered, then zinc finger polypeptides might be designed to bind any chosen DNA sequence. The complex of such a polypeptide and its recognition sequence might be utilized to modulate (up or down) the transcriptional activity of the gene containing this sequence.

Zinc finger proteins have also been reported which bind to RNA. Clemens, *et al.*, (*Science*, 260:530, 1993) found that fingers 4 to 7 of TFIIIA contribute 95% of the free energy of TFIIIA binding to 5S rRNA, whereas fingers 1 to 3 make a similar contribution in binding the promoter of the 5S gene. Comparison of the two known 5S RNA binding proteins, TFIIIA and p43, reveals few homologies other than the consensus zinc ligands (C and H), hydrophobic amino acids and a threonine-tryptophan-threonine triplet motif in finger 6.

In order to redesign zinc fingers, new selective strategies must be developed and additional information on the structural basis of sequence-specific nucleotide recognition

is required. Current protein engineering efforts utilize design strategies based on sequence and/or structural analogy. While such a strategy may be sufficient for the transfer of motifs, it limits the ability to produce novel nucleotide binding motifs not known in nature. Indeed, the redesign of zinc fingers utilizing an analogy based strategy has met with only modest success (Desjarlais and Berg, *Proteins*, 12:101, 1992).

As a consequence, there exists a need for new strategies for designing additional zinc fingers with specific recognition sites as well as novel zinc fingers for enhancing or repressing gene expression. The present invention fulfills this need.